Discriminator optimal transport

Akinori Tanaka (RIKEN AIP/iTHEMS, Keio Univ.)

@NeurIPS2019, Vancouver, Canada.

Motivation 1

- Generative Adversarial Networks (Goodfellow et al. 2014) is a well known DL scheme for generating images.

- Even if we succeed in training, there are two problems:
  1. Usually, the discriminator after the training is discarded. It sounds wasteful.
  2. The generated images sometimes include unwanted structures.

- The problem 2 may be relaxed by solving the problem 1, i.e. by recycling the trained discriminator.

Motivation 2

There are some known works in the same spirit:

- Discriminator Rejection Sampling (Azadi et al. 2018)

The ideal discriminator $D(x) = \dfrac{1}{1 + \left(\frac{p_G(x)}{p(x)}\right)} \Rightarrow$ **Rejection sampling for** $G(z)$

- Metropolis-Hastings GAN (Turner et al. 2018)

The ideal discriminator $D(x) = \dfrac{1}{1 + \left(\frac{p_G(x)}{p(x)}\right)} \Rightarrow$ **Metropolis-Hastings test for** $G(z_1) \to G(z_2) \to \ldots$

SLOGAN: "Reject if D(G(z)) is too small."

These are **passive** methods to improve generated images.

Is is possible to take an **"active"** method, somehow ?

Our proposal

- Discriminator Optimal Transport (DOT)

SLOGAN: "Deform G(z) so that D(G(z)) is large."

Target space DOT

$$\textbf{argmin}_x \left\{ ||x - G(z_y)||_2 - D(x) \right\}$$

Latent space DOT

$$\textbf{argmin}_z \left\{ ||z - z_y||_2 - D(G(z)) \right\}$$

✔ These have theoretical background from GAN's objective and OT theory.
✔ These can be implemented by gradient descent, i.e. backprop, and GPU friendly.
✔ These do improve scores like EMD, IS and FID.

For more detail, please come to our poster!
We are looking forward to discuss with you!